



## Bringing errors into focus

Nicolas Bouleau

### ► To cite this version:

| Nicolas Bouleau. Bringing errors into focus. 2007. hal-00144572

**HAL Id: hal-00144572**

**<https://hal.science/hal-00144572>**

Preprint submitted on 3 May 2007

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# BRINGING ERRORS INTO FOCUS

NICOLAS BOULEAU\*

*Lecture at the colloquium of mathematics of the University of Strasbourg, France  
7th April 2006*

## *Summary*

- I. The dichotomy of small errors, p 2,
- II. Intrinsic error calculi, p 4,
- III. Complete and symmetric error calculus, p 13,
- IV. The four bias operators, p 21,
- V. Statistics and errors, p 26.

hal-00144572, version 1 - 3 May 2007

---

\*Ecole des Ponts, ENPC, ParisTech, 28 rue des Saints Pères, 75007 Paris; e-mail : [bouleau@enpc.fr](mailto:bouleau@enpc.fr)

# I. The dichotomy of small errors.

There are two kinds of small errors which do not follow the same differential calculus.

In order to understand this phenomenon, let us consider two researchers, two applied mathematicians say, who endeavor to perform random simulation rigorously. Both are expert in using inversion or rejection methods so that they are able to simulate any probability law as soon as they can draw a real number randomly in the interval  $[0, 1]$ .

For this, congruence methods are of course available, statistically excellent, but instead they want to be able to assess the committed error :

- the first one draws the binary digits by heads or tails.
- the second one uses Polya's urn.

Let us look at the biases and the variances of the errors after  $n$  drawings.

In the first case  $x = 0, a_1 a_2 a_3 \dots$  and we approach

$$x = \sum_{k=1}^{\infty} \frac{a_k}{2^k} \quad \text{by} \quad x_n = \sum_{k=1}^n \frac{a_k}{2^k} \quad a_k \in \{0, 1\}$$

The  $a_i$  are independent random variables with value 1 or 0 with probability  $\frac{1}{2}$ ,  $\mathcal{F}_n = \sigma(a_1, \dots, a_n)$ . Let us put

$$\begin{aligned} b_n &= \mathbb{E}[(x - x_n) | \mathcal{F}_n] \\ d_n &= \mathbb{E}[(x - x_n)^2 | \mathcal{F}_n] \\ v_n &= d_n - b_n^2. \end{aligned}$$

We have

$$b_n = \sum_{n+1}^{\infty} \frac{1/2}{2^k} = \frac{1}{2^{n+1}}, \quad d_n = \frac{1}{3} \frac{1}{4^n}, \quad v_n = \frac{1}{3} \frac{1}{4^{n+1}}.$$

In the second case, let us recall the rule of Polya's urn. At the beginning there are one white and one black ball in the urn and each time a ball is drawn out, it is put back into the urn together with another one of the same color.

After  $n$  drawings, the ratio of white balls in the urn may be written

$$X_{n+1}(n+3) = X_n(n+2) + 1_{\{U_{n+1} \leq X_n\}}$$

where  $U_{n+1}$  is uniformly distributed on  $[0, 1]$  independent of  $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$ .

In other words

$$X_{n+1} = X_n + \frac{1}{n+3}(1_{\{U_{n+1} \leq X_n\}} - X_n).$$

We see that  $X_n$  is a bounded martingale which converges a.s. and in  $L^p$ ,  $p \in [1, \infty[$ , to  $X_\infty$  and it is not difficult to prove that when the initial composition of the urn is one white and one black ball,  $X_\infty$  is uniformly distributed on  $[0, 1]$ .

We have for the bias and the variance

$$b_n = \mathbb{E}[X_\infty - X_n | \mathcal{F}_n] = 0$$

$$v_n = d_n = \mathbb{E}[(X_\infty - X_n)^2 | \mathcal{F}_n] \quad \text{with} \quad \mathbb{E}[v_n] = \frac{1}{6n} + o(1/n).$$

We see that in the first case the variances are smaller than the biases, and in the second case the biases are smaller than the variances.

How will this propagate through the computations of our two modelers ?

Let us write the Taylor expansion of a  $\mathcal{C}^3$ -function with bounded derivatives.

$$f(X) - f(X_n) = (X - X_n)f'(X_n) + \frac{1}{2}(X - X_n)^2 f''(X_n) + \frac{1}{6}(X - X_n)^3 f'''(X_n + \theta(X - X_n))$$

$$\begin{aligned} B_n &= \mathbb{E}[f(X_\infty) - f(X_n) | \mathcal{F}_n] = b_n f'(X_n) + \frac{1}{2} d_n f''(X_n) + o(d_n) \\ D_n &= \mathbb{E}[(f(X_\infty) - f(X_n))^2 | \mathcal{F}_n] = d_n f'^2(X_n) + o(d_n) \end{aligned}$$

Three cases appear (supposing that the derivatives  $f'$  and  $f''$  do not vanish for simplicity, or that the sets of their zeros are negligible for the law of  $X_n$ ) :

1°/ When the variance is negligible with respect to the bias,  $v_n \ll b_n$ , (first researcher), we have also  $d_n \ll b_n$ , asymptotically the dominating term for the bias  $B_n$  is the first term  $b_n f'(X_n)$ . We see that  $D_n$  is negligible before  $B_n$  so that  $V_n \ll B_n$ , the situation will endure. The only useful formula is

$$\mathbb{E}[f(X_\infty) - f(X_n) | \mathcal{F}_n] = b_n f'(X_n) + o(b_n).$$

2°/ When the variance is of the same order of magnitude as the bias, this situation will persist ( $D_n$  and  $V_n$  are of the same order of magnitude as  $B_n$ ).

3°/ When the bias is negligible before the variance  $b_n \ll v_n$  (second researcher), the dominating term of the bias becomes  $\frac{1}{2} d_n f''(X_n)$  which is equivalent to  $\frac{1}{2} v_n f''(X_n)$  and

we fall back into the second case ( $2^\circ$ ) where the bias  $B_n$  and the variance  $V_n$  are of the same order of magnitude.

In the cases  $2^\circ$  and  $3^\circ$ ,  $b_n^2$  may be neglected before  $v_n$  and the useful formulae are

$$\begin{aligned}\mathbb{E}[f(X_\infty) - f(X_n)|\mathcal{F}_n] &= b_n f'(X_n) + \frac{1}{2}v_n f''(X_n) + o(v_n) \\ \mathbb{E}[(f(X_\infty) - f(X_n))^2|\mathcal{F}_n] &= v_n f'^2(X_n) + o(v_n).\end{aligned}$$

We see that our first modeler may content himself with an error calculus based on the first derivatives.

Instead, the second modeler, who uses Polya's urn, needs an error calculus for the bias and the variances.

- the calculus for the variances is a first order calculus,
- the one for the biases is a second order calculus which uses the calculus on variances.

In practice, generally, we do not control the nature of the errors. In a modelization, errors on data are exogenous, we do not know exactly from where they come. Therefore it is wise to do as if we were in the second case, especially to take in account the stochastic nature of the errors and the non-linearity of the model.

This may be said otherwise : to consider that a sensitivity analysis on the data or on the parameters may be done with the first derivatives, is to suppose that the errors on these quantities are representable by their mean. If, for more and more accurate measurements, the errors are thought as a stochastic process, this amount to consider it satisfies the ordinary differential calculus instead of the Ito calculus. We shall come back to this analogy later on.

## II. Intrinsic error calculi.

In the error calculus on biases and variances, the calculus on variances do not involve the biases, it is natural to begin with it, and this takes us back to Gauss at the beginning of the XIX-th century.

### Gauss' error calculus on variances and covariances.

Twelve years after his demonstration that, under some hypotheses, error follow the normal law, Gauss is interested in the propagation of errors (*Theoria combinationis* 1821). Given a quantity  $U = F(X_1, X_2, \dots)$  function of other quantities  $X_1, X_2, \dots$ , he

consider the problem of computing the quadratic error on  $U$  knowing the quadratic errors  $\sigma_1^2, \sigma_2^2, \dots$  on  $X_1, X_2, \dots$ , assuming these errors are small and independent.

His answer is the following :

$$(1) \quad \sigma_U^2 = \left(\frac{\partial F}{\partial X_1}\right)^2 \sigma_1^2 + \left(\frac{\partial F}{\partial X_2}\right)^2 \sigma_2^2 + \dots$$

and for another function  $V = G(X_1, X_2, \dots)$  he gives the covariance of the errors on  $U$  and  $V$ :

$$(2) \quad cov_{UV} = \frac{\partial F}{\partial X_1} \frac{\partial G}{\partial X_1} \sigma_1^2 + \frac{\partial F}{\partial X_2} \frac{\partial G}{\partial X_2} \sigma_2^2 + \dots$$

Gauss didn't study the propagation of biases, which, as we shall see, is more delicate.

The erring ways of textbooks : "ugly" formulae.

In spite of the works of Gauss, during the whole XIX-th century and also in the XX-th century, have been taught in mathematical or physical textbooks very ambiguous formulae like

$$(3) \quad \Delta U = \left| \frac{\partial F}{\partial X_1} \right| \Delta X_1 + \left| \frac{\partial F}{\partial X_2} \right| \Delta X_2 + \dots$$

as, for example, in the course *Mathématiques générales* (1947) of Vessiot and Montel, where the  $\Delta X_1, \Delta X_2, \Delta U$ , are the absolute value of the estimated error. Sometimes (3) is justified (J. Taylor *An Introduction to Error Analysis*, University Science Books 1992) by the inequality

$$(4) \quad \sqrt{\left(\frac{\partial F}{\partial X_1}\right)^2 \sigma_1^2 + \dots + \left(\frac{\partial F}{\partial X_k}\right)^2 \sigma_k^2} \leq \left| \frac{\partial F}{\partial X_1} \right| \sigma_1 + \dots + \left| \frac{\partial F}{\partial X_k} \right| \sigma_k.$$

which would make Gauss' formulae (1) and (2) useless. In addition to the fact that in (3) it is not at all clear what are  $\Delta X_1, \Delta X_2, \Delta U$ , these formulae are *clumsy*.

It is not a question of taste or aesthetics, it is an affair of symbolism and concepts : with (3),  $|\Delta U|$  generally depends on the way the function  $F$  is written. By composing two linear maps with values in  $\mathbb{R}^2$  we see that *the identity map increases the errors ... impossible to work properly in such conditions.*

This does not happen in the calculus of Gauss. To see this, we may introduce the differential operator

$$L = \frac{1}{2} \sigma_1^2 \frac{\partial^2}{\partial X_1^2} + \frac{1}{2} \sigma_2^2 \frac{\partial^2}{\partial X_2^2} + \dots$$

and remark that (1) writes

$$\sigma_U^2 = LF^2 - 2FLF.$$

Then the coherence is related to the transport of a differential operator by a function. If  $u$  and  $v$  are regular injective mappings and if we denote  $\theta_u L$  the operator  $\varphi \mapsto L(\varphi \circ u) \circ u^{-1}$ , we have  $\theta_{v \circ u} L = \theta_v(\theta_u L)$ . We shall make this more precise later on.

### Geometrization.

Errors on  $X_1, X_2, \dots$  may be supposed non-independent and may depend on the values of  $X_1, X_2, \dots$ : we consider a field of positive symmetric matrices  $\sigma_{ij}(x_1, x_2, \dots)$  on  $\mathbb{R}^d$  representing the conditional variances and covariances of the errors given the values  $x_1, x_2, \dots$  of  $X_1, X_2, \dots$ , and the error calculus becomes

$$(5) \quad \sigma_F^2 = \sum_{ij} \frac{\partial F}{\partial X_i}(x_1, x_2, \dots) \frac{\partial F}{\partial X_j}(x_1, x_2, \dots) \sigma_{ij}(x_1, x_2, \dots)$$

In order to geometrize the error calculus, i.e. to find a language for the biases and the variances which depends on the only mathematical objects and not on their written form, we proceed in two steps.

First we argue on  $\mathbb{R}^d$ , setting about it in such a way that if

$$F = f \circ g = k \circ h$$

the error depends only on  $F$ . Then we shall extend this when  $F$  takes its values in a manifold.

**a)** Let  $F$  be a regular function,  $\mathcal{C}^\infty$  say, of two variables  $x$  and  $y$  say. Let us consider the increment of  $F$  between  $(x, y)$  and  $(x + dx, y + dy)$ . Here  $dx$  and  $dy$  are arbitrary increments.

Taylor formula gives

$$\Delta F = F(x+dx, y+dy) - F(x, y) = P_1(dx, dy) + \dots + \frac{1}{n!} P_n(dx, dy) + (|dx| + |dy|)^n o(|dx| + |dy|)$$

where

$$P_k(dx, dy) = \sum_{p+q=k} \frac{k!}{p!q!} \frac{\partial^k F}{\partial^p x \partial^q y} dx^p dy^q$$

is a homogeneous polynomial of degree  $k$  in  $dx, dy$  with coefficients depending on  $x, y$  which we shall call the differential of order  $k$  of  $F$  denoted  $d^k F$ .

From the formula

$$\Delta F = dF + \frac{1}{2}d^2 F + \cdots + \frac{1}{n!}d^n F + (|dx| + |dy|)^n o(|dx| + |dy|)$$

we deduce the differential calculus of interest for us : if  $x = f(u, v, w)$  and  $y = g(u, v, w)$  and if  $Z(u, v, w) = F(f(u, v, w), g(u, v, w))$  we have

$$(6) \quad \left\{ \begin{array}{l} dZ = P_1(du, dv, dw) = \frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy \\ d^2 Z = P_2(du, dv, dw) = d^2 F \quad (\text{quadratic terms in } dx, dy) \\ \quad + \frac{\partial F}{\partial x}d^2 x + \frac{\partial F}{\partial y}d^2 y \\ \quad (\text{linear terms in } d^2 x, d^2 y, \text{ quadratic in } du, dv, dw) \\ \\ = \frac{\partial^2 F}{\partial x^2}(dx)^2 + 2\frac{\partial^2 F}{\partial y \partial x}dxdy + \frac{\partial^2 F}{\partial y^2}(dy)^2 + \frac{\partial F}{\partial x}d^2 x + \frac{\partial F}{\partial y}d^2 y \end{array} \right.$$

what may be remembered by writing

$$\begin{aligned} d^2 Z = d(dF) &= d\left(\frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy\right) \\ &= d\left(\frac{\partial F}{\partial x}\right)dx + \frac{\partial F}{\partial x}d^2 x + d\left(\frac{\partial F}{\partial y}\right)dy + \frac{\partial F}{\partial y}d^2 y. \end{aligned}$$

**b) Randomization.** Formulae (6) are identically valid when the increments  $du, dv, dw$  are supposed to be random.

Simply the Landau symbols  $o(\cdot)$  are now random and we must be carefull to the integrability when expectations are taken.

Thus, we may write with capitals as is usual in probability theory : if  $Z = F(X, Y)$

$$(7) \quad \left\{ \begin{array}{l} dZ = \frac{\partial F}{\partial x}dX + \frac{\partial F}{\partial y}dY \\ d^2 Z = \frac{\partial^2 F}{\partial x^2}(dX)^2 + 2\frac{\partial^2 F}{\partial y \partial x}dXdY + \frac{\partial^2 F}{\partial y^2}(dY)^2 + \frac{\partial F}{\partial x}d^2 X + \frac{\partial F}{\partial y}d^2 Y. \end{array} \right.$$

If the functions  $f, g$  and  $F$  have bounded derivatives and if the random variables  $dU, dV, dW$  are square integrable, we see that  $dX, dY$  and  $dZ$  are square integrable and  $d^2 X, d^2 Y$  and  $d^2 Z$  are in  $L^1$ .



We introduce now the bias and variance operators  $A$  and  $\Gamma$  associated to the erroneous random variable  $(X, Y)$  :

$$(8) \quad \begin{cases} A[F](x, y) = \mathbb{E}[dZ + \frac{1}{2}d^2Z \mid X=x, Y=y] \\ \Gamma[F](x, y) = \mathbb{E}[(dZ)^2 \mid X=x, Y=y] \end{cases}$$

Let us point out at once that in the formula giving  $A[F]$  we may or not put  $dZ$ , we shall come back on this, we let it for the moment.

- Let us remark that if we apply  $A$  to the function  $F^2$ , we obtain  $d^2(F^2) = d(2FdF) = 2(dF)^2 + 2Fd^2F$ , so that

$$(9) \quad \begin{aligned} A[F^2] &= \mathbb{E}^{x,y}[2FdF + (dF)^2 + Fd^2F] = 2FA[F] + \Gamma[F] \\ \Gamma[F] &= A[F^2] - 2FA[F]. \end{aligned}$$

- If we consider

$$H(X, Y) = \Phi(F_1(X, Y), F_2(X, Y), \dots, F_d(X, Y))$$

we see by the change of variable formulae (7) that

$$(10) \quad \begin{cases} A[H] = \sum_{i=1}^d \frac{\partial \Phi}{\partial F_i} A[F_i] + \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2 \Phi}{\partial F_i \partial F_j} \Gamma[F_i, F_j] \\ \Gamma[H] = \sum_{i,j=1}^d \frac{\partial \Phi}{\partial F_j} \frac{\partial \Phi}{\partial F_i} \Gamma[F_i, F_j] \end{cases}$$

the second relation generalizes that of Gauss to the case of possibly correlated errors.

**c) The analogy of formulae (10) with Ito's formula for semi-martingales and their brackets is striking. It may be made more precise as follows :**

Let us consider the vector field

$$\begin{aligned} A[X](x, y) &= \mathbb{E}[dX + \frac{1}{2}d^2X \mid X=x, Y=y] \\ A[Y](x, y) &= \mathbb{E}[dY + \frac{1}{2}d^2Y \mid X=x, Y=y] \end{aligned}$$

and the field of matrices

$$\underline{\Gamma}(x, y) = \begin{pmatrix} \Gamma[X](x, y) & \Gamma[X, Y](x, y) \\ \Gamma[X, Y](x, y) & \Gamma[Y](x, y) \end{pmatrix}$$

and let  $\Sigma(x, y)$  be a regular square root of  $\underline{\Gamma}$ , let us consider the diffusion process  $S$  solution to the stochastic differential equation

$$\begin{cases} dS_t^1 = \Sigma_{11}(S_t)dB_t^1 + \Sigma_{12}(S_t)dB_t^2 + A[X](S_t)dt \\ dS_t^2 = \Sigma_{21}(S_t)dB_t^1 + \Sigma_{22}(S_t)dB_t^2 + A[Y](S_t)dt \end{cases}$$

where  $B = (B^1, B^2)$  is a standard Brownian motion. Ito's calculus gives

$$dF_i(S)_t = \frac{\partial F_i}{\partial X}(S_t)dS_t^1 + \frac{\partial F_i}{\partial Y}(S_t)dS_t^2 + \frac{1}{2}[\frac{\partial^2 F_i}{\partial X^2}(S_t)\Gamma[X](S_t) + 2\frac{\partial^2 F_i}{\partial X \partial Y}(S_t)\Gamma[X, Y](S_t) + \frac{\partial^2 F_i}{\partial Y^2}(S_t)\Gamma[Y](S_t)]dt$$

and also, denoting  $(.)^*$  the continuous finite variation part

$$(11) \left\{ \begin{array}{l} (dF_i(S)_t)^* = \frac{\partial F_i}{\partial X}(S_t)A[X](S_t)dt + \frac{\partial F_i}{\partial Y}(S_t)A[Y](S_t)dt \\ \quad + \frac{1}{2}[\frac{\partial^2 F_i}{\partial X^2}(S_t)\Gamma[X](S_t) + 2\frac{\partial^2 F_i}{\partial X \partial Y}(S_t)\Gamma[X, Y](S_t) + \frac{\partial^2 F_i}{\partial Y^2}(S_t)\Gamma[Y](S_t)]dt \\ \text{and} \\ d\langle F_i(S), F_j(S) \rangle_t = \frac{\partial F_i}{\partial X}(S_t)\frac{\partial F_j}{\partial X}(S_t)\Gamma[X](S_t)dt + (\frac{\partial F_i}{\partial X}\frac{\partial F_j}{\partial Y} \\ \quad + \frac{\partial F_i}{\partial Y}\frac{\partial F_j}{\partial X})\Gamma[X, Y]dt + \frac{\partial F_i}{\partial Y}(S_t)\frac{\partial F_j}{\partial Y}(S_t)\Gamma[Y](S_t)dt \end{array} \right.$$

so that the change of variable formulae for the biases and the variances of errors (10) are fulfilled if we were defining  $A[F_i](x, y)$  and  $\Gamma[F_i, F_j](x, y)$  by the relations

$$A[F_i](S_t) = \frac{(dF_i(S)_t)^*}{dt}$$

$$\Gamma[F_i, F_j](S_t) = \frac{d\langle F_i(S), F_j(S) \rangle_t}{dt}$$

the same for  $H$  and any regular function of  $X$  and  $Y$ .

To sum up

We see that the notion of erroneous random variable  $(X, Y)$  with values in  $\mathbb{R}^2$  may be represented by

a random differential operator of order 1

$$F \mapsto b_{(X, Y)}[F] = \frac{\partial F}{\partial X}(X, Y)dX + \frac{\partial F}{\partial Y}(X, Y)dY$$

and a random differential operator of order 2

$$F \mapsto a_{(X, Y)}[F] = \frac{\partial F}{\partial x}d^2X + \frac{\partial F}{\partial y}d^2Y + \frac{\partial^2 F}{\partial x^2}(dX)^2 + 2\frac{\partial^2 F}{\partial y \partial x}dXdY + \frac{\partial^2 F}{\partial y^2}(dY)^2$$

where  $dX, dY, d^2X, d^2Y$  are a priori any random variables.

Then, for an error calculus retaining only the two first moments of the conditional law of the error given  $(X, Y)$ , we sum up by two (deterministic) differential operators the bias operator

$$A[F](x, y) = \mathbb{E}[b_{(X,Y)} + \frac{1}{2}a_{(X,Y)}[F] | (X, Y) = (x, y)]$$

and the operator of variance

$$\Gamma[F](x, y) = \mathbb{E}[(b_{(X,Y)}[F])^2 | (X, Y) = (x, y)].$$

and we have the change of variable formulae

$$(12) \quad \begin{cases} A[H] = \sum_{i=1}^d \frac{\partial \Phi}{\partial F_i} A[F_i] + \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2 \Phi}{\partial F_i \partial F_j} \Gamma[F_i, F_j] \\ \Gamma[H] = \sum_{i,j=1}^d \frac{\partial \Phi}{\partial F_j} \frac{\partial \Phi}{\partial F_i} \Gamma[F_i, F_j]. \end{cases}$$

a) Let us remark that we have identically

$$(b_{(X,Y)}[F])^2 = \frac{1}{2}a_{(X,Y)}[F^2] - F a_{(X,Y)}[F]$$

therefore

$$(13) \quad \Gamma[F] = A[F^2] - 2FA[F]$$

the operator  $\Gamma$  may be deduced from the operator  $a_{(x,y)}$ .

b) If instead of  $a_{(x,y)}$  we had taken

$$\hat{a}_{(X,Y)} = c_{(X,Y)} + a_{(X,Y)}$$

where  $c$  is a random field of first order differential operators, and so

$$\hat{A}[F] = \mathbb{E}[b_{(X,Y)}[F] + \frac{1}{2}\hat{a}_{(X,Y)}[F] | (X, Y) = (x, y)]$$

formulae (12) and (13) would be still verified with  $\hat{A}$ ,  $\Gamma$  remaining unchanged.

Remarks a) and b) above show that, for propagating the errors, if we keep the only operators  $A$  and  $\Gamma$  and formulae (12), then operator  $A$  discloses some ambiguity for its first order terms. We might forget the first order operator  $b$ . This may be understood by the very nature of the notion of bias which needs an origin for reference. This reference changes if we add at the beginning a deterministic error, of parallax type, which propagates following a first order operator.

d) In manifolds.

Let us recall that given a manifold  $M$ , a *second order tangent vector* at  $a$  is a differential operator at point  $a$ , without constant term, of order  $\leq 2$ , we denote  $\tau_a(M)$  their set and  $\tau(M)$  the space of fields of second order vectors.

A second order differential form is a  $\mathcal{C}^\infty$ -function on  $\tau(M)$  linear on each  $\tau_a(M)$ . We denote  $\tau^*(M)$  the space of second order differential forms.

Example : If  $f$  and  $g$  are real functions on  $M$  we may define the second order form  $d^2f = \lambda \in \tau(M) \mapsto \lambda(f)$  and then the second order form

$$df \cdot dg = \frac{1}{2}(d^2(fg) - fd^2g - gd^2f).$$

We say that  $\lambda \in \tau(M)$  is of elliptic type if  $\langle \lambda, df \cdot df \rangle \geq 0 \quad \forall f$ . We denote  $\tau^e(M)$  the space of fields of elliptic type second order vectors.

We consider a random variable  $X$  with values in  $M$  and, at each point  $x \in M$ , a probability measure on  $\tau_x^e(M)$  regular with respect to  $x$ .

In other words, we consider a random field of second order tangent vectors of elliptic type (only its marginal laws of order one will be used afterwards). We denote this field  $\Delta X$  : it is the "error" on  $X$ .

Then we define

$$(14) \quad \begin{cases} A[f](x) = \frac{1}{2}\mathbb{E}[\langle d^2f, \Delta X \rangle | X = x] \\ \Gamma[f](x) = \mathbb{E}[\langle df \cdot df, \Delta X \rangle | X = x] \end{cases}$$

$A$  is a field of second order vectors and  $\Gamma$  is a field of bilinear differential operators of positive type given by

$$\Gamma[f] = A[f^2] - 2fA[f].$$

If we consider  $h = \varphi(f_1, f_2, \dots, f_k)$ , denoting  $x^i$  a coordinate system, we have

$$\begin{aligned} A[\varphi(f_1, f_2, \dots, f_k)](x) &= \frac{1}{2}\mathbb{E}_x[\langle \Delta X, d^2h \rangle] \\ &= \frac{1}{2}\mathbb{E}_x[\langle \Delta X, \sum_i D_i h d^2x^i + \sum_{ij} D_{ij} h dx^i \cdot dx^j \rangle] \end{aligned}$$

$$\begin{aligned} \Gamma[\varphi(f_1, f_2, \dots, f_k)](x) &= \mathbb{E}_x[\langle \Delta X, dh \cdot dh \rangle] \\ &= \mathbb{E}_x[\langle \Delta X, \sum_{ij} D_i h D_j h dx \cdot dx^j \rangle] \end{aligned}$$

with

$$\begin{aligned} D_i h &= \sum_p \varphi'_p D_i f_p \\ D_{ij} h &= \sum_{p,q} \varphi''_{pq} D_i f_p D_j f_q + \sum_p \varphi'_p D_{ij} f_p \end{aligned}$$

thus, getting rid :

$$(15) \quad \begin{cases} A[\varphi(f_1, \dots, f_k)] = \sum_p \varphi'_p A[f_p] + \frac{1}{2} \sum_{p,q} \varphi''_{pq} \Gamma[f_p, f_q] \\ \Gamma[\varphi((f_1, \dots, f_k))] = \sum_{p,q} \varphi'_p \varphi'_q \Gamma[f_p, f_q]. \end{cases}$$

The interpretation in terms of diffusion processes is similar to that of the flat case. We know that if the Ito differential is  $dF(S)_t$ , its finite variation part  $(dF(S)_t)^*$  is intrinsic (cf. Paul-André Meyer, “Géométrie stochastique sans larmes” in *Sém de Probabilités XV*, Lectures Notes 850, Springer 1981, p 51), with the interpretation given in the flat case,  $F \mapsto A[F]$  taken in  $S_t$  is the second order tangent vector of the local characteristics of  $S$ .

Let us remark that if we add to the error  $\Delta X$  a random first order tangent vector, changing  $\Delta X$  in  $\Delta X + b$ , formula (14) becomes

$$A_b[f](x) = \frac{1}{2} \mathbb{E}_x[<df, b> + <d^2 f, \Delta X>]$$

because  $d^2 f|_{TM} = df$  (cf. P.-A. Meyer *op. cit.*, p 49) what does not change  $\Gamma$  and  $A_b$  still satisfies formulae (15).

Hence we see that the interpretation in terms of errors agrees with the fact that on a manifold, the first order part of a second order tangent vector is not defined [except if a linear connection is available to share the deterministic and the stochastic part (continuous local martingale) of the error]. Already in the flat case, the first order part of the bias needs a convention at the start.

Remark. Why did we limit ourselves retaining from the conditional law of the error only the two first moments, since the formalism of  $n$ -th order differential forms would make it possible to deal with the propagation of the other moments ? First let us say for the sake of simplicity, but a more precise answer will be given later on ( Remark in part IV page 23).

### III. Complete and symmetric error calculus.

a) The symmetry as invariant.

If we start from a situation where the error is centered, by a non-linear mapping the error is no more centered.

What is preserved by image ?

If  $X_n$  is an approximation of  $X$ , and if the joint law of the pair  $(X, X_n)$  is symmetric, then the law of  $(\psi(X), \psi(X_n))$  is of course symmetric too. Concretely this represents situations where we ignore between two close random variables  $X_0$  and  $X_1$  which is the right one which is the erroneous one, and by this hesitation we work with  $(X_B, X_{1-B})$  where  $B$  is a Bernoulli independent variable.

Let us take the notation of the flat case. If  $(X, X + \Delta X)$  is a symmetric pair, we have

$$\begin{aligned} F(X + \Delta X) - F(X) &= dF + \frac{1}{2}d^2F + \text{remainder} \\ G(X + \Delta X) - G(X) &= dG + \frac{1}{2}d^2G + \text{remainder} \end{aligned}$$

so

$$\begin{aligned} \mathbb{E}[F(X + \Delta X)G(X) - F(X)G(X + \Delta X)] \\ &= \mathbb{E}[G(X)(dF + \frac{1}{2}d^2F) - F(X)(dG + \frac{1}{2}d^2G) + \text{remainder}] \\ &= \mathbb{E}[G(X)A[F](X)] - \mathbb{E}[F(X)A[G](X)] + \text{remainder} \\ &= \langle G, AF \rangle_\nu - \langle F, AG \rangle_\nu + \text{remainder} \end{aligned}$$

denoting  $\nu$  the law of  $X$ .

We see that as soon as the pair  $(X, X + \Delta X)$  is symmetric and the errors small, the operator  $A$  is symmetric with respect to the law of  $X$ . This property of symmetry (which solves the ambiguity of the first order terms of  $A$ ) is an invariant by image : the bias operator of the error of the image of  $X$  by an application  $\varphi$  is symmetric with respect to the law of  $\varphi(X)$ .

The symmetric case allows to construct a much more powerful framework thanks to the theory of Dirichlet forms, and this framework extends easily to the infinite dimension (cf. N. Bouleau, *Error calculus for finance and physics*, De Gruyter 2003). We sketch it now :

b) Error structures

An error structure is a term

$$(\Omega, \mathcal{A}, \mathbb{P}, \mathbb{D}, \Gamma)$$

where  $(\Omega, \mathcal{A}, \mathbb{P})$  is a probability space, satisfying the following properties :

- 1.)  $\mathbb{D}$  is a dense sub-vectorspace of  $L^2(\Omega, \mathcal{A}, \mathbb{P})$ ,
- 2.)  $\Gamma$  is a symmetric bilinear map from  $\mathbb{D} \times \mathbb{D}$  into  $L^1(\mathbb{P})$  satisfying the Gauss calculus of class  $\mathcal{C}^1 \cap \text{Lip}$ ,  
i.e. if  $u \in \mathbb{D}^m$  and  $v \in \mathbb{D}^n$ , and if  $F$  and  $G$  are of class  $\mathcal{C}^1$  and Lipschitz, from  $\mathbb{R}^m$  [resp.  $\mathbb{R}^n$ ] into  $\mathbb{R}$ , then  $F \circ u \in \mathbb{D}$  and  $G \circ v \in \mathbb{D}$  and

$$\Gamma[F \circ u, G \circ v] = \sum_{i,j} F'_i(u) G'_j(v) \Gamma[u_i, v_j] \quad \mathbb{P}\text{-p.s.},$$

- 3.) the bilinear form  $\mathcal{E}[f, g] = \mathbb{E}\Gamma[f, g]$  is closed,  
i.e.  $\mathbb{D}$  is complete for the norm  $\|\cdot\|_{\mathbb{D}} = (\|\cdot\|_{L^2(\mathbb{P})}^2 + \mathcal{E}[\cdot, \cdot])^{\frac{1}{2}}$ ,
- 4.)  $1 \in \mathbb{D}$  (hence  $\Gamma[1, 1] = 0$ , markovianity).

We write  $\mathcal{E}[f]$  for  $\mathcal{E}[f, f]$  and  $\Gamma[f]$  for  $\Gamma[f, f]$ .

With this definition, the form  $\mathcal{E}$  is a *Dirichlet form* (local and admitting a square field operator).

To this form corresponds a Dirichlet operator  $A$  (generator of the semi-group associated to  $\mathcal{E}$ ) which, under suitable hypotheses on  $F$ , satisfies :

$$A[F \circ u] = \sum_i F'_i \circ u \ A[u_i] + \frac{1}{2} \sum_{i,j} F''_{ij} \circ u \ \Gamma[u_i, u_j] \quad \mathbb{P}\text{-p.s..}$$

Example 1. ( Ornstein-Uhlenbeck structure in dimension 1)

$$\begin{aligned}
\Omega &= \mathbb{R} \\
\mathcal{A} &= \text{Borel } \sigma\text{-field } \mathcal{B}(\mathbb{R}) \\
\mathbb{P} &= \mathcal{N}(0, 1) \text{ reduced normal law} \\
\mathbb{D} = H^1(\mathcal{N}(0, 1)) &= \{u \in L^2(\mathbb{P}), u' \text{ in distributions sense} \\
&\quad \text{belongs to } L^2(\mathbb{P})\} \\
\Gamma[u] &= u'^2
\end{aligned}$$

then,  $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathcal{N}(0, 1), H^1(\mathcal{N}(0, 1)), \Gamma)$  is an error structure. We can also obtain the bias operator (the associated generator) :

$$\mathcal{D}A = \{f \in L^2(\mathbb{P}): f'' - xf' \text{ in the sense of distributions} \in L^2(\mathbb{P})\}$$

and  $Af = \frac{1}{2}f'' - \frac{1}{2}I \cdot f'$  where  $I$  is the identity map on  $\mathbb{R}$ .

Example 2. ( Monte-Carlo structure in dimension 1)

$$\begin{aligned}
\Omega &= [0, 1] \\
\mathcal{A} &= \text{Borel } \sigma\text{-field} \\
\mathbb{P} &= \text{Lebesgue measure} \\
\mathbb{D} &= \{u \in L^2([0, 1], dx): \text{the derivative } u' \text{ in the sense of distributions} \\
&\quad \text{on } ]0, 1[ \text{ belongs to } L^2([0, 1], dx)\} \\
\Gamma[u] &= u'^2.
\end{aligned}$$

the space  $\mathbb{D}$  thus defined is denoted  $H^1([0, 1])$ .

Example 3. Let  $U$  be a domain (connected open set) of  $\mathbb{R}^d$  with unit volume,  $\mathcal{B}(U)$  the Borel  $\sigma$ -field and  $dx = dx_1, \dots, dx_d$  the Lebesgue measure,

$$\begin{aligned}
\mathbb{D} &= \{u \in L^2(U, dx): \text{the gradient } \nabla u \text{ in the sense of distributions} \\
&\quad \text{belongs to } L^2(U, dx; \mathbb{R}^d)\} \\
\Gamma[u] &= |\nabla u|^2 = \left(\frac{\partial u}{\partial x_1}\right)^2 + \dots + \left(\frac{\partial u}{\partial x_d}\right)^2.
\end{aligned}$$



Then  $(U, \mathcal{B}(U), dx, \mathbb{D}, \Gamma)$  is an error structure. From the relation  $\mathcal{E}[f, g] = \langle -Af, g \rangle$  it follows that the domain of the generator contains the functions of class  $\mathcal{C}^2$  with compact support in  $U$ ,  $\mathcal{D}A \supset \mathcal{C}_K^2(U)$  and that for such functions

$$Af = \frac{1}{2}\Delta f = \frac{1}{2} \sum_{i=1}^d \frac{\partial^2 f}{\partial x_i^2}.$$

Example 4.

Let  $D$  be an open set in  $\mathbb{R}^d$  with unit volume. Let  $\mathbb{P} = dx$  be the Lebesgue measure on  $D$ . Let  $\Gamma$  be defined on  $\mathcal{C}_K^\infty(D)$  by

$$\Gamma[u, v] = \sum_{ij} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} a_{ij}, \quad u, v \in \mathcal{C}_K^\infty(D)$$

where the functions  $a_{ij}$  satisfy the following hypotheses

- $a_{ij} \in L_{\text{loc}}^2(D) \quad \frac{\partial a_{ij}}{\partial x_k} \in L_{\text{loc}}^2(D) \quad i, j, k = 1, \dots, d$
- $\sum_{i,j} a_{ij}(x) \xi_i \xi_j \geq 0 \quad \forall \xi \in \mathbb{R}^d \quad \forall x \in D$
- $a_{ij}(x) = a_{ji}(x) \quad \forall x \in D.$

Then the pre-structure  $(D, \mathcal{B}(D), \mathbb{P}, \mathcal{C}_K^\infty(D), \Gamma)$  is closable.

**Comment on the completeness.**

The fact that we ask the form  $\mathcal{E}$  be *closed* is a restriction, but this restriction is highly fruitful. The situation is quite analogous to the question of the  $\sigma$ -additivity in probability theory : without this property, nothing can be said on the transmission of errors by objects defined by limits. But many objects in contemporary mathematics are defined by limits (integrals, solutions to ode, solutions to sdp, stochastic integrals, solutions to sde, etc.)<sup>1</sup>

---

<sup>1</sup>The philosopher Karl Popper is fallen in this trap by emphasizing that his own theory of probability (additive) do contain strictly the one of Kolmogorov ( $\sigma$ -additive) Cf. N.Bouleau "Some thoughts upon axiomatized languages, a focus on probability theory and error calculus with Dirichlet forms" Butlletí de la Societat Catalana de Matemàtiques Vol. 18 n2 p25-36, (2004) cf. <http://hal.archives-ouvertes.fr/hal-00105636>

This extension tool allows

- to extend the error calculus on variances from  $\mathcal{C}^1$ -functions to Lipschitz functions,
- to prove a criterion of existence of density for the probability laws which generalizes Malliavin's method : the image energy density property.

The closedness property is preserved by images and products even infinite products. This permits to construct naturally errors structures on the spaces of stochastic processes.

In particular on the Wiener space, where may be obtained, among others, the Ornstein-Uhlenbeck operator. As well, the Malliavin calculus may be interpreted as an error calculus.

### Theorem on products

Let  $S_n = (\Omega_n, \mathcal{A}_n, \mathbb{P}_n, \mathbb{D}_n \Gamma_n)$ ,  $n \geq 1$ , be error structures. The product structure

$$S = (\Omega, \mathcal{A}, \mathbb{P}, \mathbb{D}, \Gamma) = \prod_{n=1}^{\infty} S_n$$

is defined by

$$(\Omega, \mathcal{A}, \mathbb{P}) = \left( \prod_{n=1}^{\infty} \Omega_n, \otimes_{n=1}^{\infty} \mathcal{A}_n, \prod_{n=1}^{\infty} \mathbb{P}_n \right)$$

$$\begin{aligned} \mathbb{D} = \Big\{ f \in L^2(\mathbb{P}) : & \forall n, \text{ for almost every } w_1, w_2, \dots, w_{n-1}, w_{n+1}, \dots \\ & \text{for the product measure} \\ & x \mapsto f(w_1, \dots, w_{n-1}, x, w_{n+1}, \dots) \in \mathbb{D}_n \text{ and} \\ & \int \sum_n \Gamma_n[f] d\mathbb{P} < +\infty \Big\} \end{aligned}$$

and for  $f \in \mathbb{D}$

$$\Gamma[f] = \sum_{n=1}^{\infty} \Gamma_n[f].$$

$S$  is an error structure, Markovian if each  $S_n$  is.

When we write  $\Gamma_n[f]$ ,  $\Gamma_n$  acts only on the  $n$ -th argument of  $f$ .

Starting from the basic bricks that are one dimensional errors structures, by product we obtain error structures on function spaces.

This yields easily error structures on

- the Wiener space, cf. N. B. & F. Hirsch, *Dirichlet Forms and Analysis on Wiener Space*, De Gruyter 1991,
- the general Poisson space, the Monte Carlo space, cf. N. B. *Error Calculus for Finance and Physics*, De Gruyter, 2003.

## Images of error structures.

The operation is as mere and almost as general as the image of a probability measure by a measurable map.

If  $(\Omega, \mathcal{A}, \mathbb{P}, \mathbb{D}, \Gamma)$  is an error structure and  $X$  a random variable with values in  $\mathbb{R}^d$  whose components are in  $\mathbb{D}$ , the term  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mathbb{P}_X, \mathbb{D}_X, \Gamma_X)$  is an error structure, where

$$\begin{aligned}\mathbb{P}_X & \text{ is the law of } X, \\ \mathbb{D}_X & = \{f \in L^2(\mathbb{P}_X) : f \circ X \in \mathbb{D}\} \\ \Gamma_X[f](x) & = \mathbb{E}\{\Gamma[f \circ X] | X=x\}, \quad f \in \mathbb{D}.\end{aligned}$$

In fact it is possible to define images by more general random variables.

The image structure by  $X$  may be called the "Dirichlet-law" of  $X$ . It is an error structure on  $\mathbb{R}^d$  such that the identity map  $I$  has its components in the domain of  $\Gamma_X$  and the following formulae hold :

$$\begin{aligned}\mathbb{E}[\Gamma[X]|X] & = \Gamma_X[I] \circ X \\ \mathbb{E}[\Gamma[\varphi(X)]|X] & = \Gamma_X[\varphi] \circ X.\end{aligned}$$

Several theorems of probability theory possess analogs in the theory of errors structures under some conditions : Gateaux-Lévy, Strassen, etc. (cf. Bouleau-Hirsch [de Gruyter 1991].)

## The case of Wiener space.

- Let  $(\chi_n)_{n \in \mathbb{N}}$  be an orthonormal basis of  $L^2(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+), dx)$ , and let  $(g_n)_{n \in \mathbb{N}}$  be a sequence of i.i.d. reduced Gaussian random variables.

To a function  $f \in L^2(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+), dx)$  is associated the Wiener integral

$$I(f) = \sum_n \langle f, \chi_n \rangle g_n,$$

homomorphism of  $L^2(\mathbb{R}_+, dx)$  into  $L^2(\Omega, \mathcal{A}, \mathbb{P})$ . If we put

$$B(t) = \sum_n \langle 1_{[0,t]}, \chi_n \rangle g_n = \sum_n \int_0^t \chi_n(y) dy \cdot g_n$$

then  $B(t)$  is a standard Brownian motion.

Because of the case where  $f$  is a step-function,  $I(f)$  is denoted  $I(f) = \int_0^\infty f(s) dB_s$

- The preceding construction uses the product space

$$(\Omega, \mathcal{A}, \mathbb{P}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathcal{N}(0, 1))^{\mathbb{N}},$$

the  $g_n$ 's being the coordinate maps. If we put an error structure on each factor

$$(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathcal{N}(0, 1), \mathbf{d}_n, \gamma_n),$$

we obtain an error structure on  $(\Omega, \mathcal{A}, \mathbb{P})$ :

$$(\Omega, \mathcal{A}, \mathbb{P}, \mathbb{D}, \Gamma) = \prod_{n=0}^{\infty} (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathcal{N}(0, 1), \mathbf{d}_n, \gamma_n)$$

such that a random variable  $F(g_0, g_1, \dots, g_n, \dots)$  belongs to  $\mathbb{D}$  iff

$$\begin{cases} \forall n & x \mapsto F(g_0, \dots, g_{n-1}, x, g_n, \dots) \text{ belongs to } \mathbf{d}_n \quad \mathbb{P}\text{-p.s.} \\ \text{and} & \Gamma[F] = \sum_n \gamma_n[F] \text{ belongs to } L^1(\mathbb{P}) \quad (\gamma_n \text{ acting on the } n\text{-th variable of } F). \end{cases}$$

- If on each factor we take the Ornstein-Uhlenbeck structure in dimension 1, we obtain

$$\begin{aligned}\Gamma[g_n] &= 1 \\ \Gamma[g_m, g_n] &= 0 \quad \text{if } m \neq n.\end{aligned}$$

For  $f \in L^2(\mathbb{R}_+)$ , from  $\int_0^\infty f(s) dB_s = \sum_n \langle f, \chi_n \rangle g_n$  we deduce

$$\Gamma \left[ \int_0^\infty f(s) dB_s \right] = \sum_n \langle f, \chi_n \rangle^2 = \|f\|_{L^2(\mathbb{R}_+)}^2,$$

hence, using the functional calculus  $\forall F \in \mathcal{C}^1 \cap \text{Lip}(\mathbb{R}^m)$

$$\Gamma \left[ F \left( \int f_1(s) dB_s, \dots, \int f_n(s) dB_s \right) \right] = \sum_{i,j} \frac{\partial F}{\partial x_i} \frac{\partial F}{\partial x_j} \int f_i(s) f_j(s) ds.$$

It is the Ornstein-Uhlenbeck structure on the Wiener space.

- For the sake of simplicity let us restrict the time to  $t \in [0, 1]$  and let us take for  $\chi_n$  the trigonometric basis. If on each factor we take the Ornstein-Uhlenbeck structure scaled by a constant coefficient depending on  $n$  :

$$\prod_{n=0}^{\infty} (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathcal{N}(0, 1), H^1(\mathcal{N}(0, 1)), u \rightarrow (2\pi n)^{2q} u^2)$$

we obtain on the Wiener space an error structure satisfying on the first chaos :

$$\Gamma \left[ \int_0^1 f(s) dB_s \right] = \int_0^1 (f^{(q)})^2(s) ds$$

where  $f^{(q)}$  is the  $q$ -th derivative of  $f$ . It is a structure where the error disturbs longitudinally the Brownian path, which belongs to the family of structures called of "generalized Mehler type".

## IV. The four bias operators.

We tackle now the following question : How an error, in the usual sense in mathematics, i.e. an approximation error, generates an error structure.

Let us consider a random variable  $Y$  defined on the probability space  $(\Omega, \mathcal{A}, \mathbb{P})$  with values in the measurable space  $(E, \mathcal{F})$  and approximations  $Y_n$ ,  $n \in \mathbb{N}$ , also defined on  $(\Omega, \mathcal{A}, \mathbb{P})$  with values in  $(E, \mathcal{F})$ .

We suppose there exist an algebra  $\mathcal{D}$  of bounded functions from  $E$  into  $\mathbb{R}$  or  $\mathbb{C}$  dense in  $L^2(E, \mathcal{F}, \mathbb{P}_Y)$  containing the constants and a sequence  $(\alpha_n)_{n \in \mathbb{N}}$  of positive numbers, about which we consider the following hypotheses :

$$(H1) \quad \begin{cases} \forall \varphi \in \mathcal{D}, \text{ there exists } \overline{A}[\varphi] \in L^2(E, \mathcal{F}, \mathbb{P}_Y) \quad \text{s.t.} \quad \forall \chi \in \mathcal{D} \\ \lim_{n \rightarrow \infty} \alpha_n \mathbb{E}[(\varphi(Y_n) - \varphi(Y))\chi(Y)] = \mathbb{E}_Y[\overline{A}[\varphi]\chi] \end{cases}$$

the expectation  $\mathbb{E}_Y$  being relative to the law  $\mathbb{P}_Y$ .

$$(H2) \quad \begin{cases} \forall \varphi \in \mathcal{D}, \text{ there exists } \underline{A}[\varphi] \in L^2(E, \mathcal{F}, \mathbb{P}_Y) \quad \text{s.t.} \quad \forall \chi \in \mathcal{D} \\ \lim_{n \rightarrow \infty} \alpha_n \mathbb{E}[(\varphi(Y) - \varphi(Y_n))\chi(Y_n)] = \mathbb{E}_Y[\underline{A}[\varphi]\chi]. \end{cases}$$

$$(H3) \quad \begin{cases} \forall \varphi \in \mathcal{D}, \text{ there exists } \tilde{A}[\varphi] \in L^2(E, \mathcal{F}, \mathbb{P}_Y) \quad \text{s.t.} \quad \forall \chi \in \mathcal{D} \\ \lim_{n \rightarrow \infty} \alpha_n \mathbb{E}[(\varphi(Y_n) - \varphi(Y))(\chi(Y_n) - \chi(Y))] = -2\mathbb{E}_Y[\tilde{A}[\varphi]\chi]. \end{cases}$$

As soon as two hypotheses among (H1) (H2) (H3) are fulfilled (with the same algebra  $\mathcal{D}$  and the same sequence  $\alpha_n$ ), the third one follows thanks to the relation

$$\tilde{A} = \frac{\overline{A} + \underline{A}}{2}.$$

- When defined the operator  $\overline{A}$  which considers the asymptotic error from the point of view of the limit model, will be called *the theoretical bias operator*.

- The operator  $\underline{A}$  which consider the error from the point of view of the approximate model, will be called *the practical bias operator*.

- Because of the property

$$\langle \tilde{A}[\varphi], \chi \rangle_{L^2(\mathbb{P}_Y)} = \langle \varphi, \tilde{A}[\chi] \rangle_{L^2(\mathbb{P}_Y)}$$

the operator  $\tilde{A}$  will be called *the symmetric bias operator*.

The following result shows that a Dirichlet form (possibly non-local) is often present behind an approximation :

**Theorem.** *Under hypothesis (H3)*

*a) the limit*

$$\tilde{\mathcal{E}}[\varphi, \chi] = \lim_n \frac{\alpha_n}{2} \mathbb{E}[(\varphi(Y_n) - \varphi(Y))(\chi(Y_n) - \chi(Y))] \quad \varphi, \chi \in \mathcal{D}$$

*defines a closable bilinear form whose smallest closed extension is denoted  $(\mathcal{E}, \mathbb{D})$ .*

*b)  $(\mathcal{E}, \mathbb{D})$  is a Dirichlet form*

*c)  $(\mathcal{E}, \mathbb{D})$  admits a square of field operator  $\Gamma$  satisfying  $\forall \varphi, \chi \in \mathcal{D}$*

$$\Gamma[\varphi] = \tilde{A}[\varphi^2] - 2\varphi\tilde{A}[\varphi]$$

$$\mathbb{E}_Y[\Gamma[\varphi]\chi] = \lim_n \alpha_n \mathbb{E}[(\varphi(Y_n) - \varphi(Y))^2(\chi(Y_n) + \chi(Y))/2]$$

*d)  $(\mathcal{E}, \mathbb{D})$  is local if and only if  $\forall \varphi \in \mathcal{D}$*

$$\lim_n \alpha_n \mathbb{E}[(\varphi(Y_n) - \varphi(Y))^4] = 0.$$

We introduce now the fourth bias operator  $\mathbb{A}$  defined under (H1) and (H2) on  $\mathcal{D}$  by

$$\mathbb{A} = \frac{1}{2}(\overline{A} - \underline{A}).$$

Since  $\mathbb{E}_Y[\mathbb{A}[\varphi]\chi] = \lim_n \mathbb{E}[(\varphi(Y_n) - \varphi(Y))(\chi(Y) + \chi(Y_n))/2]$  we see that  $\mathbb{A}$  represents the asymptotic error from the point of view of an outside observer who attaches the same weight both to the theoretical and the practical models and measuring the error algebraically on the same axis. Because of properties we shall state below, the operator  $\mathbb{A}$  will be called *the singular bias operator*.

An operator  $B$  from  $\mathcal{D}$  into  $L^2(\mathbb{P}_Y)$  is said *first order* if it satisfies

$$B[\varphi\chi] = B[\varphi]\chi + \varphi B[\chi] \quad \forall \varphi, \chi \in \mathcal{D}.$$

**Proposition.** Under (H1) to (H3)

a) the theoretical variance  $\lim_n \alpha_n \mathbb{E}[(\varphi(Y_n) - \varphi(Y))^2 \psi(Y)]$  and the practical variance  $\lim_n \alpha_n \mathbb{E}[(\varphi(Y_n) - \varphi(Y))^2 \psi(Y_n)]$  exist and we have  $\forall \varphi, \chi, \psi \in \mathcal{D}$

$$\begin{aligned} \lim_n \alpha_n \mathbb{E}[(\varphi(Y_n) - \varphi(Y))(\chi(Y_n) - \chi(Y))\psi(Y)] &= \mathbb{E}_Y[-\underline{A}[\varphi\psi]\chi + \underline{A}[\psi]\varphi\chi - \overline{A}[\varphi]\chi\psi] \\ \lim_n \alpha_n \mathbb{E}[(\varphi(Y_n) - \varphi(Y))(\chi(Y_n) - \chi(Y))\psi(Y_n)] &= \mathbb{E}_Y[-\overline{A}[\varphi\psi]\chi + \overline{A}[\psi]\varphi\chi - \underline{A}[\varphi]\chi\psi] \end{aligned}$$

b) These two variances coincide if and only if  $\mathbb{A}$  is first order, and then they are equal to  $\mathbb{E}_Y[\Gamma[\varphi]\psi]$ .

c) If the Dirichlet form is local, then  $\mathbb{A}$  is first order.

**Remark.** Under (H3) condition d) of the theorem (p 22) characterising the case where the form  $\mathcal{E}$  is local is equivalent to either one of the following conditions :

$$(j) \exists \lambda > 2 \quad \lim_n \alpha_n \mathbb{E}[|\varphi(Y_n) - \varphi(Y)|^\lambda] = 0 \quad \forall \varphi \in \mathcal{D}.$$

$$(jj) \forall \lambda > 2 \quad \lim_n \alpha_n \mathbb{E}[|\varphi(Y_n) - \varphi(Y)|^\lambda] = 0 \quad \forall \varphi \in \mathcal{D}.$$

This yields an answer to the question put on page 12 about an error calculus with higher moments using tangent vectors of order greater than 2 : In the cases where the Dirichlet form is local, (the only case in which we are able to propagate errors thanks to a differential calculus) the moments of order greater than 2 are negligible before the first or second order moments.

Let us come back to the case where only hypothesis (H3) is assumed. The following result shows that, for the variances, the error calculus performed on the limit model with the asymptotic error, coincides with the error asymptotically obtained on  $\mathcal{C}^1$ -functions.

**Proposition.** Under (H3). If the form  $(\mathcal{E}, \mathbb{D})$  is local, then the asymptotic error calculus principle is valid on

$$\tilde{\mathcal{D}} = \{F(f_1, \dots, f_p) : f_i \in \mathcal{D}, F \in \mathcal{C}^1(\mathbb{R}^p, \mathbb{R})\}$$

$$\begin{aligned} i.e. \lim_n \alpha_n \mathbb{E}[(F(f_1(Y_n), \dots, f_p(Y_n)) - F(f_1(Y), \dots, f_p(Y)))^2] \\ = \mathbb{E}_Y[\sum_{i,j=1}^p F'_i(f_1, \dots, f_p) F'_j(f_1, \dots, f_p) \Gamma[f_i, f_j]]. \end{aligned}$$

### Examples

- Let us take for  $Y$  a Brownian motion  $B$  indexed by  $[0, 1]$  as random variable with value in  $\mathcal{C}([0, 1])$  and let us take for  $Y_\varepsilon$  the approximation  $Y_\varepsilon = B + \sqrt{\varepsilon}W$  where  $W$  is



an independent Brownian motion. We may apply the theorem taking for  $\mathcal{D}$  the linear combinations of functions  $\varphi(B) = e^{i \int_0^1 f dB}$  with regular  $f$  say  $\mathcal{C}_b^1$ .

Hypothesis (H3) is fulfilled. The theorem yields the Ornstein-Uhlenbeck structure on the Wiener space.

- Series with independent increments. let be

$$S = \sum_{n=1}^{\infty} \frac{X_n}{n^2} + \frac{Z_n}{n}$$

where  $X_n, Z_n \in L^{2+\varepsilon}$ ,  $Z_n$  centered,  $(X_n, Z_n)$  i.i.d., on and  $S$  is approximated by the partial sum  $S_n = \sum_{k=1}^n \frac{X_k}{k^2} + \frac{Z_k}{k}$ .

By Burkholder's inequality, we have  $n\mathbb{E}[|S - S_n|^{2+\varepsilon}] \rightarrow 0$  as  $n \rightarrow \infty$ . So that, taking  $\mathcal{D} = \mathcal{C}_K^\infty$ , for  $\varphi, \chi \in \mathcal{D}$

$$\begin{aligned} \lim_n n\mathbb{E}[(\varphi(S) - \varphi(S_n))\chi(S_n)] &= \lim_n n\mathbb{E}[(S - S_n)\varphi'(S_n)\chi(S_n) + \frac{1}{2}(S - S_n)^2\varphi''(S_n)\chi(S_n)] \\ &= \frac{1}{2}\mathbb{E}[Z_1^2]\mathbb{E}[\varphi''(S)\chi(S)] + \mathbb{E}[X_1]\mathbb{E}[\varphi'(S)\chi(S)]. \end{aligned}$$

Hence hypothesis (H2) is satisfied and

$$\underline{A}[\varphi] = \frac{\mathbb{E}[Z_1^2]}{2}\varphi'' + \mathbb{E}[X_1]\varphi'.$$

Similarly

$$\lim_n n\mathbb{E}[(\varphi(S) - \varphi(S_n))^2] = \lim_n n\mathbb{E}[(S - S_n)^2\varphi'^2(S_n)] = \mathbb{E}[Z_1^2]\mathbb{E}[\varphi'^2(S)]$$

(H3) is fulfilled as soon as the law of  $S$  satisfies the Hamza condition (cf. M. Fukushima et al. [1994]) and then the Dirichlet form is local, and  $\Gamma[\varphi] = \underline{A}[\varphi^2] - 2\varphi\underline{A}[\varphi] = \mathbb{E}[Z_1^2]\varphi'^2$ .

- Stochastic integral. Let us consider the integral

$$Y = \int_0^1 H_s dB_s$$

approximated by the sum

$$Y_n = \sum_{k=0}^{n-1} H_{\frac{k}{n}}(B_{\frac{k+1}{n}} - B_{\frac{k}{n}})$$

$(B_t)$  is a standard Brownian motion defined as coordinate process on the space  $\mathcal{C}([0, 1])$  equipped with Wiener measure,  $H_s = H_0 + \int_0^s \xi_u dB_u + \int_0^s \eta_u du$  is an Ito process regular in Malliavin's sense. Under suitable hypotheses we obtain :

- Hypothesis (H3) is fulfilled under rather simple regularity assumptions and

$$\langle \tilde{A}[\varphi], \chi \rangle = -\frac{1}{4}\mathbb{E}\left[\int_0^1 \xi_s^2 ds \varphi'(Y) \chi'(Y)\right].$$

- Hypothesis (H1) supposes finer regularity conditions and

$$\begin{aligned} \langle \overline{A}[\varphi], \chi \rangle &= \frac{1}{2}\mathbb{E}\left[\int_0^1 \xi_s D_s D_s [\varphi'(Y) \chi(Y)] ds\right] + \frac{1}{2}\mathbb{E}\left[\int_0^1 \eta_s D_s [\varphi'(Y) \chi(Y)] ds\right] \\ &\quad - \frac{1}{4}\mathbb{E}\left[\int_0^1 \xi_s^2 (\varphi' \chi)'(Y) ds\right] - \frac{1}{4}\mathbb{E}\left[\int_0^1 \xi_s^2 \varphi'(Y) \chi'(Y) ds\right] \end{aligned}$$

where  $D$  denotes Malliavin's gradient on the Wiener space. Then  $\mathbb{A}$  is a first order operator.

• Stochastic differential equations and the Euler scheme. The approximation error of the solution of an SDE by the Euler scheme has been the subject of many works whose one of the achievements is a form of functional central limit theorem. See especially Jean Jacod and Philip Protter "Asymptotic error distributions for the Euler method for stochastic differential equations" *Ann. Probab.* 26, 267-307, (1998) and the quoted references. This central limit theorem allows to simplify the study of the limit

$$\lim_{n \rightarrow \infty} \alpha_n \mathbb{E}[(\varphi(Y_n) - \varphi(Y))^2]$$

but the existence of the operator  $\tilde{A}$ , and therefore the closability of the form are yet only proved in dimension 1. On the other hand the operator  $\overline{A}$  has been determined by Paul Malliavin and Anton Thalmaier "Numerical error for SDE: asymptotic expansion and hyperdistributions" *Note C. R. A. S. SI*, vol 336, n10, p851, 2003, its appearance is similar to the one given above in the case of a stochastic integral.

## V. Statistics and errors.

In the error calculus by Dirichlet forms *any erroneous quantity is random*. The a priori probability law may be thought as "the field and the use" of the measurement device.

How to determine by experiment the operator  $\Gamma$  of an error structure ?

Let us suppose that the error structure to be identified be on  $\mathbb{R}^d$ :

$$(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mathbb{P}, \mathbb{D}, \Gamma).$$

Concretely  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mathbb{P})$  is the image space of a quantity  $x$  which is measured with some accuracy.  $\mathbb{P}$  is its a priori law.

We shall consider that to carry out a measurement of the quantity  $x$ , is to estimate  $x$  statistically as parameter of a family of probability measures  $\mathbb{Q}_x$ .

It is known that if  $T$  is an estimate of  $x$  unbiased say  $(\mathbb{Q}_x[T] = x)$ , the precision on  $x$  is limited by the inequality

$$\mathbb{Q}_x[(T - x)(T - x)^t] \geq [J(x)]^{-1}$$

with equality if  $T$  is efficient (Cramer-Darmois-Fisher-Rao inequality). Let us recall the definition of the Fisher information and the Cramer-*et-al.* inequality.

Let be  $x \in \mathbb{R}^d$  and let  $\mathbb{Q}_x$  be a family of probability measures on some space dominated by the probability measure  $\mathbb{Q}$

$$\mathbb{Q}_x = L(x, \cdot) \mathbb{Q} \quad \text{with } L(x, \cdot) \text{ regular in } x.$$

Then for every random variable  $Y \in L^2(\mathbb{Q})$

$$\mathbb{E}_x[Y - \mathbb{E}_x(Y)]^2 \geq (\text{grad} \mathbb{E}_x(Y))^t [J(x)]^{-1} \text{grad} \mathbb{E}_x(Y)$$

where  $J(x)$  is the Fisher information matrix of the model

$$J(x) = \left( \mathbb{E}_x \left[ \frac{\partial \log L(x)}{\partial x_i} \frac{\partial \log L(x)}{\partial x_j} \right] \right)_{ij}.$$

Let  $I$  be the identity map from  $\mathbb{R}^d$  into itself,  $J(x)$  behaves as an information,  $\Gamma[I](x)$  is a precision.

We assume the identification :

$$\boxed{\Gamma[I](x) = J^{-1}(x)}$$

If we choose this way  $\Gamma[I](x) = J^{-1}(x)$ , since  $\Gamma$  satisfies the functional calculus, if  $f : \mathbb{R}^d \longrightarrow \mathbb{R}^p$  is of class  $\mathcal{C}^1 \cap \text{Lip}$ , we have

$$\Gamma[f](x) = (\text{grad} f)^t \cdot \Gamma[I](x) \cdot \text{grad} f$$

In other words the operator  $\Gamma$  is uniquely determined, and similarly for all image structures of our error structure  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mathbb{P}, \mathbb{D}, \Gamma)$ .

Hence the question arises to know whether we obtain a precision compatible with this calculus when for  $f$  one to one we measure  $f(x)$ . The answer is positive.

Under the hypotheses of a "regular" statistical model, if we consider  $y = f(x)$  for  $f \in \mathcal{C}^1 \cap \text{Lip}$  and one to one, the error structure obtained for  $y$  is the image of the error structure obtained for  $x$  by the map  $f$ . The identification is also stable by products in a natural sense. Cf. N. B. and Chr. Chorro, "Structures d'erreur et estimation paramétrique", *Note C.R.A.S.*, Ser.I 338 (2004), 305-310.

That means that the error on  $x$  obtained this way *do not depend on parametrization* and *possesses a physical meaning*.

## Some books related to these subjects

AZEMA, J.; YOR, M. *Séminaire de Probabilités XVI, 1980/81 Supplément : Géométrie Différentielle Stochastique*, LNM 921, Springer, 1982.

BOULEAU N. *Error Calculus for Finance and Physics, the Language of Dirichlet Forms*, De Gruyter, 2003.

BOULEAU N., HIRSCH F. *Dirichlet Forms and Analysis on Wiener Space*, De Gruyter, (1991).

DELLACHERIE, CL., MEYER, P.-A., *Probabilités et Potentiel*, Hermann 1987.

EMERY, M. *Stochastic Calculus in Manifolds*, Springer, 1989.

FUKUSHIMA, M.; OSHIMA, Y.; TAKEDA, M. *Dirichlet Forms and Symmetric Markov Processes*, De Gruyter 1994.

JACOD, J., SHIRYAEV, A.N., *Limit Theorems for Stochastic Processes*, Springer 1987.

MA, Z.-M., RÖCKNER, M. *Introduction to the Theory of (Non-symmetric) Dirichlet Forms*, Springer 1992.

MALLIAVIN, P., THALMAIER, A. *Stochastic Calculus of Variations in Mathematical Finance*, Springer 2005.

NUALART, N. : *The Malliavin Calculus and Related Topics*, Springer, 1995.